Supplemental Material

Table S1. Pairwise sequence identities for GA and GB variants

	GA^{wt}	GB^{wt}	GA^{30}	GB^{30}	GA^{77}	GB^{77}	GA^{88}	GB^{88}	GA^{95}	GB^{95}
GA^{wt}	-	16%	95%	25%	61%	38%	57%	45%	54%	48%
GB^{wt}	9	-	21%	91%	52%	75%	55%	67%	59%	64%
GA^{30}	53	12	-	30%	66%	43%	63%	50%	59%	54%
GB^{30}	14	51	17	-	61%	84%	64%	77%	68%	73%
GA^{77}	34	29	37	34	-	77%	96%	84%	93%	88%
GB^{77}	21	42	24	47	43	-	80%	93%	84%	89%
GA^{88}	32	31	35	36	54	45	-	88%	96%	91%
GB^{88}	25	38	28	43	47	52	49	-	91%	93%
GA^{95}	30	33	33	38	52	47	54	51	-	95%
GB^{95}	27	36	30	41	49	50	51	52	53	-

Table S2. Statistics of CS-Rosetta/CS23D models for GA^{wt} and GB^{wt} and variants $GA^{88/95}$ and $GB^{88/95}$.

	CS-Rosetta [a]		*CS-Rosetta (δ^1H only)* [b]		*CS23D*	
	$RMSD_{bb}$ [c]	$RMSD_{all}$ [d]	$RMSD_{bb}$ [c]	$RMSD_{all}$ [d]	$RMSD_{bb}$ [c]	$RMSD_{all}$ [d]
GA^{wt}	1.42±0.11	2.11±0.12	1.33±0.10	2.08±0.07	4.88±0.35	6.29±0.30
GB^{wt}	0.58±0.12	1.32±0.14	0.88±0.18	1.56±0.22	0.24±0.03	0.48±0.01
GA^{88}	1.91±0.25	2.94±0.21	1.95±0.20	2.80±0.19	9.93±0.12	10.90±0.13
GB^{88}	1.25±0.12	2.23±0.19	1.31±0.06	2.07±0.13	1.26±0.04	2.43±0.03
GA^{95}	1.78±0.12	2.89±0.11	1.81±0.14	2.77±0.15	9.15±0.13	10.62±0.12
GB^{95}	1.45±0.21	2.47±0.23	1.58±0.13	2.38±0.18	1.42±0.04	2.44±0.04

[a] results for CS-Rosetta runs with chemical shifts for backbone and $^{13}C^{\beta}$ atoms.

[b] results for CS-Rosetta runs with only ^{1}H chemical shifts.

[c] rmsd (Cα, C', and N) of the 10 lowest-energy models to the experimental NMR structure (PDB entries 2fs1, 1pga, 2jws, 2jwu, 2kdl and 2kdm for GA^{wt}, GB^{wt}, GB^{88}, GB^{88}, GA^{95} and GB^{95}, respectively). Disordered residues 1 to 8 and 54 to 56 are excluded from the rmsd calculation of GA variants.

[d] rmsd (all non-H atoms) of the 10 lowest-energy models to the experimental structure. Disordered residues 1 to 8 and 54 to 56 are excluded from the rmsd calculation of GA variants.

Table S3. Sequence identities of GA and GB variants relative to two homologues in Rosetta database.[a]

	2j5y[a] (3α)		2qmt[a] (4β+α)	
	$N^{fragment\ b}$	Identity[%] (e-score)	$N^{fragment\ b}$	Identity [%] (e-score)
GA^{95}	*6(6)/3(3)*	*42 (0.07)*	52(47)/31(31)	57 (9.3e-6)
GB^{95}	*5(5)/5(5)*	*41 (0.17)*	56(47)/31(31)	62 (4.6e-7)
GA^{88}	12(9)/37(35)	44 (0.02)	30(30)/23(23)	53 (2.9e-5)
GB^{88}	*4(4)/5(4)*	*37 (1.76)*	54(48)/31(31)	66 (1.75e-7)
GA^{77}	19(17)/37(35)	48 (6.29e-3)	20(20)/24(24)	50 (6.19e-4)
GB^{77}	*3(3)/6(5)*	*32 (31.3)*	54(47)/31(31)	73 (2.87e-9)
GA^{30}	46(39)/45(43)	51 (9.39e-4)	*8(8)/10(10)*	*16*
GB^{30}	*5(5)/6(5)*	*12*	56(50)/41(40)	89 (4.07e-9)
GA^{wt}	43(38)/45(43)	53 (2.45e-5)	*6(6)/6(6)*	*12*
GB^{wt}	*3(3)/6(5)*	*8*	57(52)/48(46)	98 (3.94e-12)

[a] These two homologues were excluded for all CS-Rosetta and standard Rosetta evaluations in the present study, but this table represents the impact on fragment selection when they are not excluded. PDB IDs of the two proteins in the Rosetta/CS-Rosetta database are 2j5y (for GA) and2qmt (for GB).

[b] Numbers of 3-residue/9-residue fragments selected by Rosetta from 2j5y or 2qmt, respectively; the numbers of target fragments in the query protein that have fragments selected from the same position in proteins 2qmt and 2j5y are listed in the parenthesis.

Figure S1. Secondary structure prediction for GA and GB variants. During a standard Rosetta fragment selection procedure, programs PsiPred, SAM-T99 and JUFO were first used to make secondary structure prediction from the sequence alone; the predicted secondary structure information was then combined and used to guide the Rosetta fragment selection. For each of the GA and GB variants, the predicted secondary structure by each program is plotted, as well as the secondary structure observed in the experimental structures of GA^{wt} and GB^{wt}.

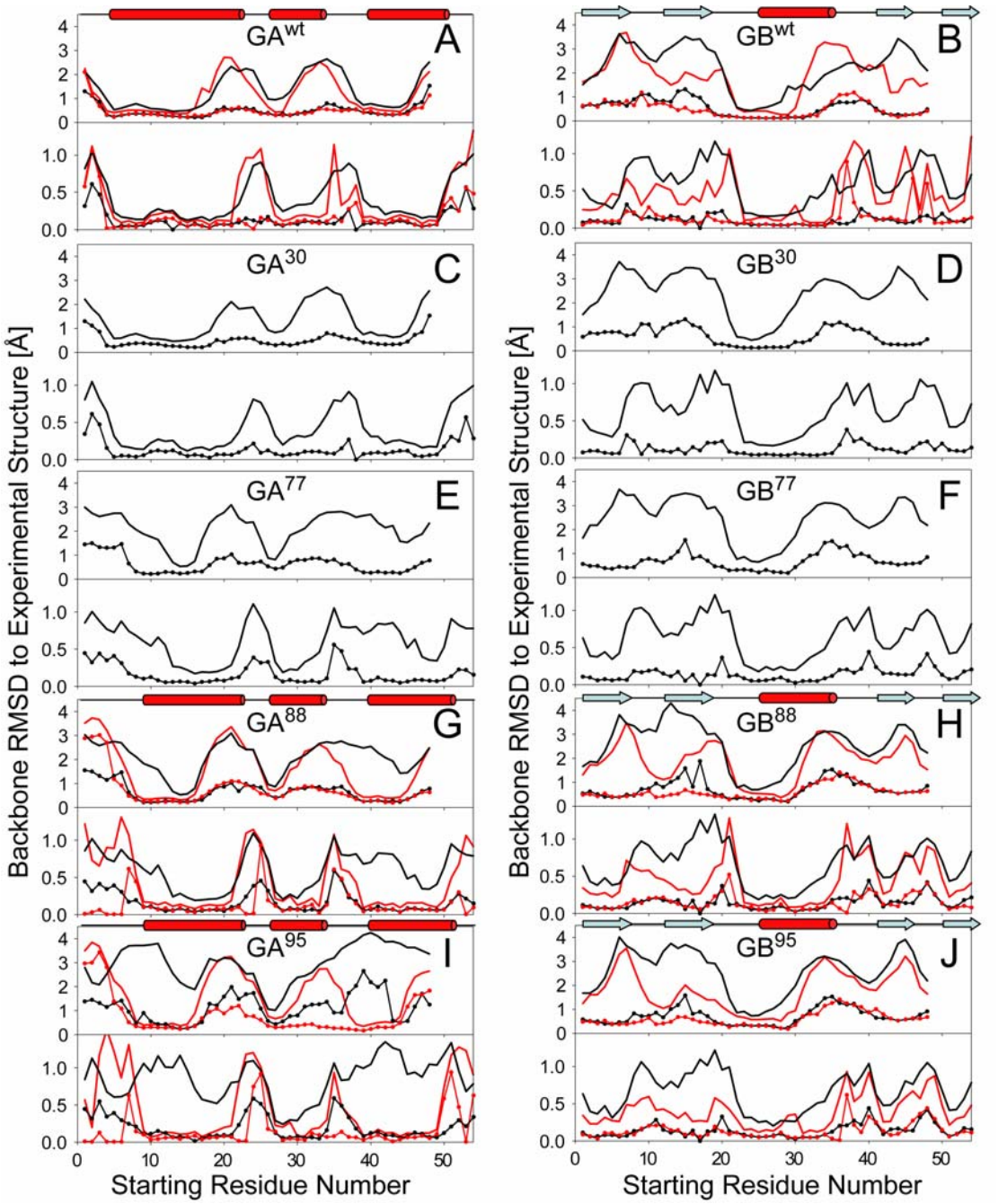

Figure S2. Quality of Rosetta/CS-Rosetta fragments for GA and GB variants, shown as plots of the lowest (lines with dots) and average (bold lines) backbone coordinate rmsd's (N, C^{α} and C') between query segment and sets of 200 9-residue (upper panel) or 3-residue (lower panel) fragments, as a function of starting position of the query segment in the sequence. Results from the standard Rosetta fragment selection method are plotted in black, while those selected using the standard MFR method with only $^{1}H^{N}$ and $^{1}H^{\alpha}$ chemical shifts are displayed in red. (A) GA^{wt}; (B) GB^{wt}; (C) GA^{30}; (D) GB^{30}; (E) GA^{77}; (F) GB^{77}; (G) GA^{88}; (H) GB^{88}; (I) GA^{95}; (J) GB^{95}. Note that no chemical shifts are available for GA^{30}, GB^{30}, GA^{77}, and GB^{77}. For 9-residue fragments, the last residue starting number in the 56-residue protein is 48; for 3-residue fragments the last starting position is 54.

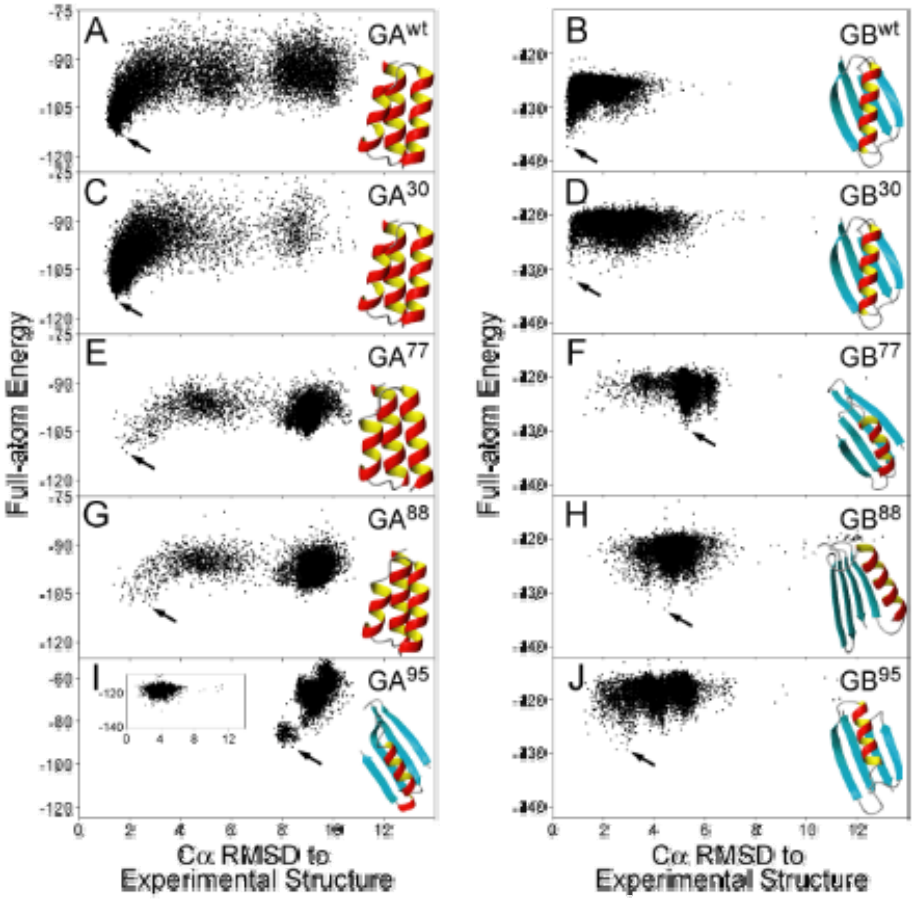

Figure S3. Energies for 10,000 models generated by standard Rosetta for GA^{wt} and GB^{wt} and their variants (while excluding the two homologous proteins of Table S3 from the Rosetta structural database). (A-J) Plots of Rosetta all-atom energy *versus* C^{α} rmsd relative to the experimental structure. (A) GA^{wt}; (B) GB^{wt}; (C) GA^{30}; (D) GB^{30}; (E) GA^{77}; (F) GB^{77}; (G) GA^{88}; (H) GB^{88}; (I) GA^{95}; (J) GB^{95}. For models of GA^{wt} and GA variants, residues 1-8 and 54-56, which exhibit a high degree of local disorder in the preliminary evaluation of lowest energy structures, are excluded from the calculation of the C^{α} rmsd and the final Rosetta energy score. For GA^{30}/GB^{30} and GA^{77}/GB^{77} Rosetta models, the C^{α} rmsd are calculated relative to the GA^{wt}/GB^{wt} and GA^{88}/GB^{88} experimental structures, respectively. The distribution of Rosetta all-atom energy of GA^{95} models *versus* C^{α} rmsd (of all residues) relative to the experimental structure of GB^{95} is plotted as an inset in (I), showing the similarity in pattern to those shown in (H) and (J) for GB^{88}/GB^{95}. The backbone ribbon representation of the lowest-energy Rosetta model, marked by an arrow, is plotted at the lower right corner of each panel.

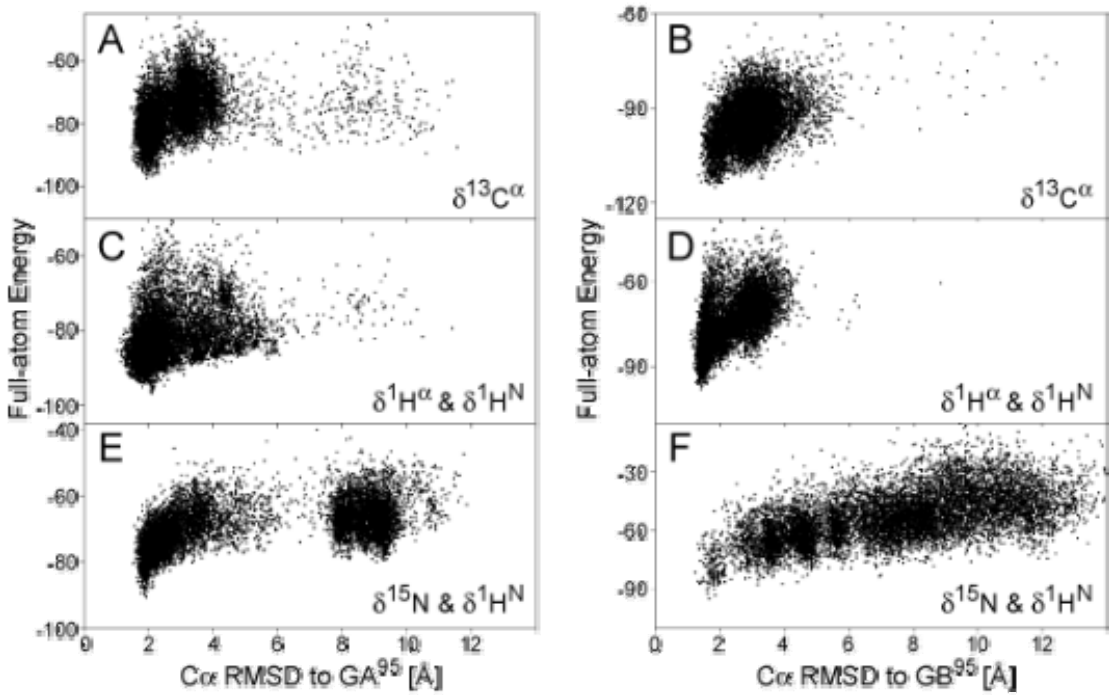

Figure S4. Energies for 10,000 models generated by CS-Rosetta for GA^{95} and GB^{95} using subsets of the NMR chemical shift assignments. (A-F) Plots of Rosetta all-atom energy, rescored by using the (partial) experimental chemical shifts, *versus* C^{α} rmsd relative to the experimental structures of GA^{95} and GB^{95}. For GA^{95}, terminal residues 1-8 and 54-56, which exhibit a high degree of local disorder in the preliminary evaluation of lowest energy structures, are excluded from the C^{α} rmsd and the calculation of the final Rosetta energy. (A and B) CS-Rosetta models for GA^{95} (A) and GB^{95} (B) obtained using only $^{13}C^{\alpha}$ chemical shifts. (C and D) CS-Rosetta models for GA^{95} (C) and GB^{95} (D) obtained using $^{1}H^{\alpha}/^{1}H^{N}$ chemical shifts. (E and F) CS-Rosetta models for GA^{95} (E) and GB^{95} (F) obtained using $^{15}N/^{1}H^{N}$ chemical shifts.

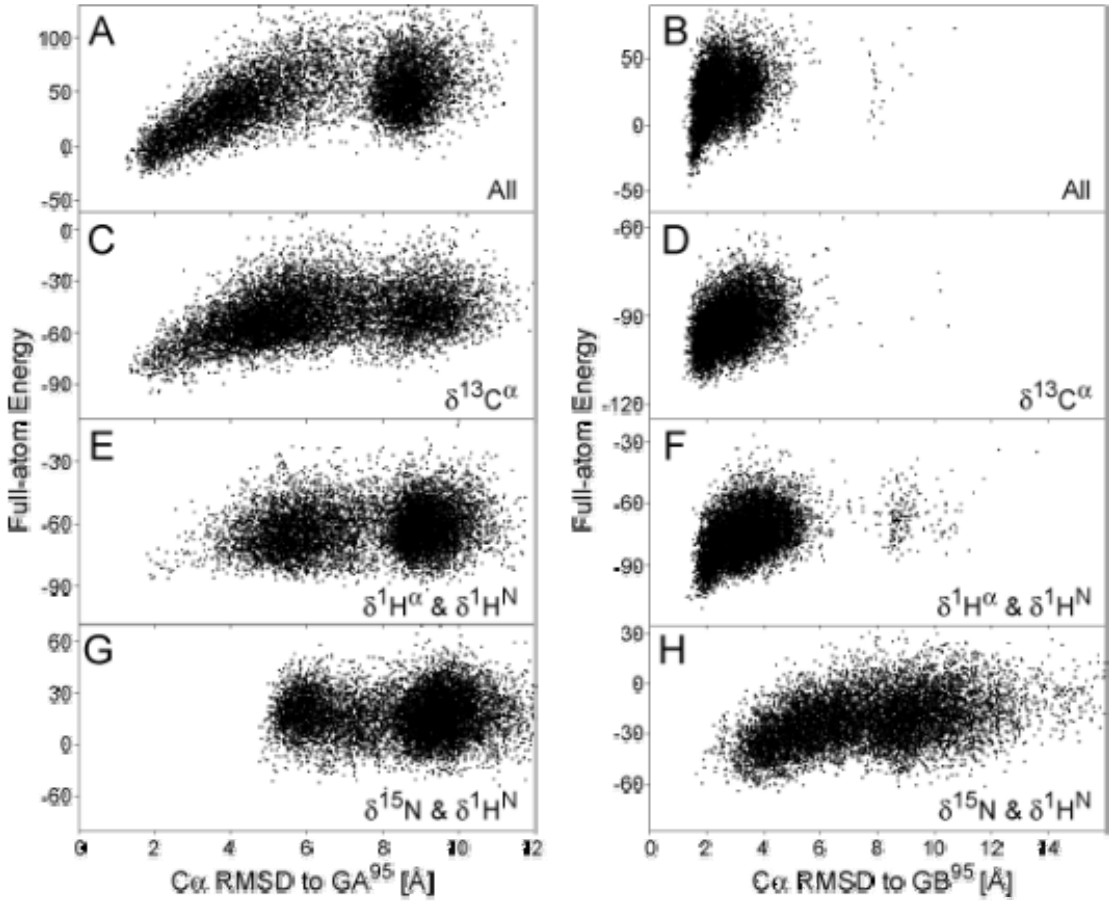

Figure S5. Energies for 10,000 models generated by CS-Rosetta for GA^{95} and GB^{95} when using hybrid fragment selection and subsets of the NMR chemical shift assignments. (A-F) Plots of Rosetta all-atom energy rescored by using the experimental chemical shifts *versus* C^{α} rmsd relative to the experimental structures for GA^{95} and GB^{95}. For GA^{95}, terminal residues 1-8 and 54-56, which exhibit a high degree of local disorder in the preliminary evaluation of lowest energy structures, are excluded from the C^{α} rmsd and the calculation of the final Rosetta energy. (A and B) CS-Rosetta models for GA^{95} (A) and GB^{95} (B) obtained using all chemical shifts. (C and D) CS-Rosetta models for GA^{95} (C) and GB^{95} (D) obtained using only $^{13}C^{\alpha}$ chemical shifts. (E and F) CS-Rosetta models for GA^{95} (E) and GB^{95} (F) obtained using $^{15}N/^{1}H^{N}$ chemical shifts.

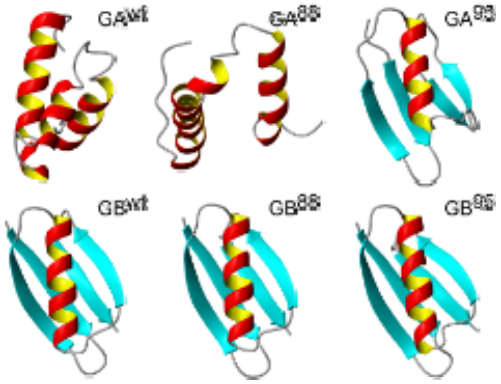

Figure S6. CS23D structure prediction for GA and GB variants. For proteins GA^{wt}, GB^{wt}, GA^{88}, GB^{88}, GA^{95}, and GB^{95}, the backbone ribbon representation of the lowest-energy CS23D models is displayed.